

# EXHIBIT A

**UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK**

THE NEW YORK TIMES COMPANY

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI, INC.,  
OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,  
OPENAI OPCO LLC, OPENAI GLOBAL LLC,  
OAI CORPORATION, LLC, and OPENAI  
HOLDINGS, LLC,

Defendants.

Civil Action No. 1:23-cv-11195-SHS

**THE NEW YORK TIMES COMPANY’S SECOND SET OF REQUESTS FOR  
PRODUCTION TO OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,  
OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, AND  
OPENAI HOLDINGS, LLC**

Pursuant to the Federal Rules of Civil Procedure and the Local Rules of the Southern District of New York, (the “Local Civil Rules”), The New York Times Company (“The Times”) requests that OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO LLC, OpenAI Global LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC (“OpenAI Defendants”) produce the requested documents at the offices of Susman Godfrey L.L.P., One Manhattan West, New York, New York 10001-8602, within thirty days after the date of service of this document, or at such other mutually agreeable time and place.

**I. DEFINITIONS**

The following definitions and instructions apply to the requests listed below:

1. All definitions and rules of construction set forth in Rule 26.3 of the Local Rules of the Southern District of New York are incorporated by reference herein.

2. “Affiliate” means, with respect to any Person, any other Person that, directly or indirectly through one or more intermediaries, controls, is controlled by, or is under any common control with, such Person.

3. “Agreement(s)” means any contract, agreement, arrangement, or understanding, formal or informal, oral or written, between or among two or more Persons.

4. “AI Model(s)” means a component of an information system and/or large language model that implements artificial intelligence (“AI”) technology or uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs.

5. “Generative AI Models(s)” means AI Models that emulate the structure and characteristics of input data in order to generate derived synthetic content, such as images, videos, audio, text, and other digital content.

6. “Text Generation AI Model(s)” means Defendants’ Generative AI Models, including generative pre-trained transformers or large language models (LLMs), that provide text outputs in response to inputs or prompts, including but not limited to GPT base, GPT-1, GPT-2, GPT-3, GPT 3.5, GPT-4, GPT-4 Turbo, GPT-4o, MAI-1, and all other versions or other Generative AI Models included in Defendants’ Generative AI Products and Services.

7. “Complaint” means the operative complaint in the Action as of the date these Requests are served.

8. “Defendants” means Microsoft and OpenAI.

9. “Electronically Stored Information” or “ESI” refers to information and Documents within the full scope of Federal Rule of Civil Procedure 34 with all Metadata intact—created, manipulated, communicated, stored, and best utilized in digital form, and stored on electronic media.

10. “Generative AI Product(s) and Service(s)” means any publicly available or commercial product or service offering that includes a Generative AI Model, including but not limited to ChatGPT branded products and services, Copilot branded products and services, ChatGPT, Browse with Bing, Bing Chat, OpenAI’s soon-to-be announced search product,<sup>1</sup> APIs, and platforms otherwise providing access to a Generative AI Model.

11. “Including” means including but not limited to.

12. “Journalism,” for the purposes of these requests, means the activity of writing or creating content for newspapers, magazines, news websites, mobile applications, television, podcasts, or any other publication and/or news outlet, and includes the work of The Times as alleged in the Complaint.

13. “Metadata” refers to information about an electronic file that is embedded in the file, describing the characteristics, origins, usage, and validity of the electronic file, and includes but is not limited to descriptive, structural, administrative, and relationship information.

14. “OpenAI” or “You” or “Your” means Defendants OpenAI Inc., OpenAI GP, LLC, OpenAI Opco, LLC, OpenAI Global, LLC, OpenAI Holdings LLC, OpenAI, LLC, or OAI Corporation, LLC, their subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

15. “Relating to” means discussing, describing, referring to, pertaining to, containing, analyzing, studying, reporting on, commenting on, evidencing, constituting, setting forth, considering, recommending, concerning, or pertaining to, in whole or in part.

---

<sup>1</sup> See Anna Tong, *OpenAI Plans to Announce Google Search Competitor on Monday, Sources Say*, REUTERS (May 10, 2024), <https://www.reuters.com/technology/openai-plans-announce-google-search-competitor-monday-sources-say-2024-05-09/>.

16. “Microsoft” means Defendant Microsoft Corporation, its subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

17. “The Times” or “The New York Times” means Plaintiff The New York Times Company, its subsidiaries, affiliates, divisions, predecessors in interest, successors, or assignees, and their respective officers, directors, employees, consultants, representatives, and agents.

18. “Times Content” means material published by The Times, including in print, online, audio, or in any other format, whether on Times-owned platforms or available through third-party sources. “Times Content” includes all New York Times interest-specific publications, sub-brands, and products, including The Athletic, Cooking, Games, and Wirecutter.

19. “Training Dataset(s)” means the data and content used to pre-train, train, fine-tune, or otherwise affect the performance of the Generative AI Models, including any data and content used for reinforcement learning from human feedback or reinforcement learning from AI feedback.

## **II. INSTRUCTIONS**

1. Produce all responsive documents prepared, sent, or received, in whole or in part, during the time period applicable to the request. You should identify in some reasonable manner the person from whose files each document you produce was found.

2. You are to produce all documents, as defined above, that are in the possession, custody or control of you, or in the possession, custody or control of any attorney or agent for you. Without limiting the terms, “possession, custody, or control” of an item means that you either have physical possession of the item or have a right to possession of the item that is equal or superior to the person who has physical possession of the item.

3. You are under a duty to supplement your responses to these discovery requests pursuant to FRCP 26. If after you produce any of the discovery requested below you discover or

receive additional documents or information responsive to the requests, you should promptly produce such additional documents or information to the full extent required by the rules.

4. Pursuant to FRCP 26, you are to provide a privilege log for any documents you have withheld on the basis of privilege(s). Request for such a log is hereby made, effective as of the response date of these requests. If you assert a claim of privilege with respect to any document request, or portion thereof, you should produce a log, accounting for each document withheld individually, and providing the nature of the privilege being claimed; the type of document, date, author, addressees or recipients, and general subject matter of the document, and such other information as is sufficient to identify the document. If a claim of privilege applies to only a portion of a document, the document should be produced, and the portion claimed to be privileged obscured and stamped “redacted.”

5. The documents requested below should be produced in the manner that they are kept in the usual course of business. Documents should be produced with the label or labels from any file folder, binder, file drawer, file box, notebook, computer disk, or other container in which the document was found. Documents on computers or computer disks, including but not limited to electronic mail and corresponding attachments, should be produced in both native format and single-page TIFF format with optical character recognition (OCR). Even if only part of a document is responsive to a request, you should still produce the entire document including all attachments.

6. All duplicates or copies of documents are to be provided to the extent they have handwriting, additions, or deletions of any kind different from the original document being produced.

7. All documents which cannot be legibly copied should be produced in their original form.

8. If any documents requested herein have been lost, discarded or destroyed, they shall be identified as completely as possible, including, without limitation, the following information: (a) the date of disposal; (b) the manner of disposal; (c) the reason for disposal; (d) the person authorizing disposal; and (e) the person disposing of the document.

9. If no Documents responsive to a particular request exist, state that no responsive Documents exist.

### **III. DOCUMENT REQUESTS**

#### **REQUEST FOR PRODUCTION NO. 16:**

Documents concerning the decision not to release Training Datasets or other information for Defendants' Text Generation AI Model(s) after GPT-3 (including all versions thereof, such as GPT-4 Turbo and GPT-4o), including any communications regarding this decision and concerns regarding the effects of releasing this data on safety, ethics, profitability, competition, and legal obligations and/or ramifications.

#### **REQUEST FOR PRODUCTION NO. 17:**

Documents concerning the compilation, acquisition, and curation of Training Datasets, including: (i) communications regarding paying for or obtaining permission to use any Training Datasets or content therein, ii) communications with Training Dataset creators, including Common Crawl, and iii) communications concerning the relevance and impact of paywalls on building Training Datasets or gathering training data.

#### **REQUEST FOR PRODUCTION NO. 18:**

Documents concerning the removal of any content, including information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use of works, from Training Datasets before, during, or after training of Defendants' Generative AI Model(s).

**REQUEST FOR PRODUCTION NO. 19:**

Documents concerning all processing performed on the Training Datasets containing Times Content or Journalism content used to train Defendants' Generative AI Products and Services, including: (i) the ordering of training samples containing Times or Journalism Content in the training process as compared to training samples not containing Times or Journalism Content; (ii) placement of training samples containing Times or Journalism Content in the training process; (iii) rules for choosing the order of training samples containing Times or Journalism Content in the training process; (iv) rules and procedures for training samples containing Times or Journalism Content that are not used on all other training samples; (v) the processing of Times or Journalism Content in the fine-tuning/instruction tuning phase of training; (vi) use of Times or Journalism Content to train or inform certain, modules, layers, mixtures of experts, or other portions of the Text Generation AI Models more than others, (vii) any measures taken to include or exclude certain portions of training samples containing Times or Journalism Content during training, and (viii) the importance of training on accurate and/or well-written content, including Times or Journalism Content.

**REQUEST FOR PRODUCTION NO. 20:**

Documents sufficient to show the total number of records and tokens in Defendants' Training Datasets and the total number of records and tokens that contain (i) Times Content, (ii) Journalism content, and (iii) copyrighted content.

**REQUEST FOR PRODUCTION NO. 21:**

Documents sufficient to show each source of records in Defendants' Training Datasets, and for each source: (i) the total number of tokens, (ii) the total number of records, (iii) the total number of tokens that contain Times Content, (iv) the total number of records that contain Times Content, (v) the total number of tokens that contain Journalism content, (vi) the total number of records that



contain Journalism content; (vii) the total number of tokens that contain copyrighted content, (viii) the total number of records that contain copyrighted content and (ix) any agreements relating to Defendants' use, licensing, or purchasing of the source of records.

**REQUEST FOR PRODUCTION NO. 22:**

Documents sufficient to identify each article of Times Content contained in Defendants' Training Datasets.

**REQUEST FOR PRODUCTION NO. 23:**

Documents concerning all steps taken or contemplated to train, instruct, program, or refine Defendants' Generative AI Products and Services using Times Content or Journalism content, including: (i) the use of Times or Journalism Content in instruction/fine tuning; (ii) use of Times Content or Journalism content in Reinforcement Learning from Human Feedback ("RLHF") or Reinforcement Learning from AI Feedback ("RLAIF"), (iii) use of Times or Journalism Content to train the reward component or function of RLHF or RLAIF, (iv) use of Times or Journalism Content to evaluate a feature of a Training Dataset, such as classifying Journalism content or identifying quality of a training sample, and (v) RLHF moderator and ranking guidance, training materials and ranking factors, weighting or scoring factors, and results/output of the human annotators, moderators and reviewers.

**REQUEST FOR PRODUCTION NO. 24:**

Documents concerning any analyses, studies, measurements, testing, experimentation, assessments, or other evaluation of the incremental value of including Journalism content, including Times Content, in Defendants' Training Datasets.

**REQUEST FOR PRODUCTION NO. 25:**

Documents concerning any analyses, studies, measurements, testing, experimentation, assessments, or other evaluation of the quantity, ordering, placement, or other configurable parameter of Journalism content, including Times Content, in Defendants' Training Datasets.

**REQUEST FOR PRODUCTION NO. 26:**

Documents concerning Your use or knowledge of web crawlers, bots, spiders, user agents, and related tools to access The Times's Content, including Times websites and digital products, including documents sufficient to show: (i) the time periods in which those tools accessed The Times's Content, websites and digital products, (ii) the Times Content those tools accessed, fetched, or parsed from The Times's websites and digital products, and (iii) the Times Content from these URLs that was crawled, processed, copied, stored, and/or indexed by You and/or systems under Your control.

**REQUEST FOR PRODUCTION NO. 27:**

Documents concerning Your reliance on, use of, or understanding of third-party robots.txt crawler directives or third-party meta tag crawler directives in managing the crawling and indexing activity of your web crawlers, bots, spiders, user agents, GPTBot, any plugins or custom GPTs included on Defendants' platforms, and related applications to access, crawl, and/or index Times Content or third-party content.

**REQUEST FOR PRODUCTION NO. 28:**

Documents concerning all measures taken to conceal or obscure Defendants' use of Times Content during any stage of the development of each of Defendants' Generative AI Products and Services, including: (i) any methods, procedures, or instructions used to prevent Defendants' Generative AI Products and Services from reproducing Training Datasets containing Times Content, (ii) any methods, procedures, or instructions used to determine if the output of

Defendants' Generative AI Products and Services contains Times Content, (iii) any methods, procedures, or instructions to avoid overfitting or memorization of Times Content in Defendants' Generative AI Models, (iv) any differences in the foregoing among Defendants' Generative AI Products and Services, and (v) any changes in the foregoing over time.

**REQUEST FOR PRODUCTION NO. 29:**

Documents concerning Defendants' reliance on earlier versions of Defendants' Generative AI Models to be used in or to train or develop later versions of Defendants' Generative AI Models, including Defendants' use of outputs, weights, or any learnings acquired from earlier versions of Defendants' Generative AI Models.

**REQUEST FOR PRODUCTION NO. 30:**

The full model cards, data set cards, or system cards for each of Defendants' Generative AI Models, including those not released to the public.

**REQUEST FOR PRODUCTION NO. 31:**

Documents relating to Your technical documentation for Defendants' Generative AI Models, Products, and Services as defined in Annex IV of the European Union's Artificial Intelligence Act.

**REQUEST FOR PRODUCTION NO. 32:**

Documents concerning the training curricula, stages, iterations, cycles, batches, or epochs for Defendants' Generative AI Models.

**REQUEST FOR PRODUCTION NO. 33:**

Documents concerning Defendants' decisions regarding which works to use to train and develop Defendants' Generative AI Models, and how to use those works in training, including (i) documents concerning the factors considered in determining the weighting or frequency at which

works are seen during the training of Defendants' models, (ii) documents concerning de-duplication of Training Datasets; (iii) documents concerning any classifiers and/or filters applied to Training Datasets that determine document quality and/or remove low quality documents; (iv) documents concerning rules and procedures for determining sample frequency; (v) documents sufficient to show how the training weights or parameters within Defendants' Generative AI Models would change if Times Content was not used during training; and (vi) documents sufficient to show which individuals made or make decisions regarding which works to use to train and develop Defendants' Generative AI Models, and how to use those works in training.

**REQUEST FOR PRODUCTION NO. 34:**

Documents sufficient to describe each aspect or phase of the process for training Defendants' Generative AI Models, including the pre-training, training, fine-tuning, and any other refinement or development processes as well as any distinctions or differences between and among the aspects or phases.

**REQUEST FOR PRODUCTION NO. 35:**

Documents concerning any changes made to Defendants' Generative AI Models or Generative AI Products and Services from 2021 to present, including any logs showing changes, documents relating to ablation or "unlearning," (i.e. the ability to remove memorization, content, or other effects of specific parts of the Training Datasets from the models), and documents and communications concerning any changes to the models made or considered, including in response to claims of copyright infringement or improper use of data.

**REQUEST FOR PRODUCTION NO. 36:**

Documents concerning the differences between Defendants' Generative AI Models, including (i) differences in the Training Datasets and/or the composition of the Training Datasets

used for each model; (ii) differences in the training curricula used for each model; and (iii) the differences between the “Turbo” and non-Turbo versions of Defendants’ models.

**REQUEST FOR PRODUCTION NO. 37:**

The following documents concerning each of Defendants’ Generative AI Models, Products, and Services: (i) product briefs, (ii) readme files, (iii) product requirement documents, (iv) product or project management tools and records (e.g., Jira), (v) documentation from workflow management programs, (vi) minutes or notes from technical team meetings, (vii) documentation concerning data loaders, (viii) documents related to red-teaming and/or stress testing, (ix) technical system or architecture documents, specifications, or reports, and (x) launch documents.

**REQUEST FOR PRODUCTION NO. 38:**

Source code related to (i) Defendants’ Generative AI Models, Products, and Services, (ii) training curricula for Defendants’ Generative AI Models, (ii) any ranking, weighting, or scoring of training data, (iii) data loaders, (iv) Your web crawlers, spiders, and bots, including GPTBot, and (v) any processing of training data.

**REQUEST FOR PRODUCTION NO. 39:**

Documents concerning the study, testing, or evaluation of Defendants’ Generative AI Products and Services and Generative AI Models for hallucinations, memorization, and any problematic behaviors, including: (i) the results of the efforts of safety teams, including any red-teaming efforts, (ii) responses to ChatGPT’s feedback contest, (iii) submissions to Your bug bounty program, (iv) any other red-teaming, and (v) the extent to which rates of hallucination and memorization have changed over time within Defendants’ Generative AI Models.

**REQUEST FOR PRODUCTION NO. 40:**

Documents concerning the prevalence of outputs that are verbatim copies of text included in Training Datasets, including any documents concerning (i) steps taken to prevent or encourage

outputs that are verbatim copies of text included in training data, (ii) the use of filter detectors, refusal messages, and other safety features used by Defendants to prevent unlawful access to copyrighted content contained within or referenced by Defendants' Generative AI Models, and (iii) documents concerning the rate at which any such safety features are triggered in response to user queries.

**REQUEST FOR PRODUCTION NO. 41:**

Documents or surveys related to users' trust in the accuracy or veracity of information contained in outputs of Defendants' Generative AI Products and Services and how users respond to disclaimers regarding the accuracy or veracity of information contained in such outputs.

**REQUEST FOR PRODUCTION NO. 42:**

Documents concerning Defendants' efforts and abilities to monitor, log, analyze, quantify, assess, evaluate, or review user queries within any Generative AI Products and Services.

**REQUEST FOR PRODUCTION NO. 43:**

Any surveys, studies, analyses, or statistics on patterns and practices any of Defendants' Generative AI Products and Services, including any mentions of The New York Times, such surveys, studies, analyses, or statistics regarding of user engagement with in user queries and/or inclusion of links to Times Content in user queries or outputs.

**REQUEST FOR PRODUCTION NO. 44:**

Documents concerning any attempts to use Defendants' Generative AI Products and Services to circumvent paywalls or other access control measures.

**REQUEST FOR PRODUCTION NO. 45:**

Query, session, and chat logs reflecting or analyzing user sessions related to Times Content within Defendants' Generative AI Products and Services, including, for each session, user queries paired with responses to those queries.

**REQUEST FOR PRODUCTION NO. 46:**

Documents sufficient to show how Defendants' Generative AI Products and Services generate suggested follow-up queries for users.

**REQUEST FOR PRODUCTION NO. 47:**

Documents concerning user engagement metrics, events, and analytics data collected during each user session, including duration, the number of prompts or queries that a user inputs per session, click-through data, fallback rates, and bounce rates.

**REQUEST FOR PRODUCTION NO. 48:**

Documents sufficient to show (i) the number of daily active users of each of Defendants' Generative AI Products and Services, separated into the numbers of new users and repeat users on each day, from 2021 to present, and (ii) dates of subscription and changes in subscriber numbers over time.

**REQUEST FOR PRODUCTION NO. 49:**

Documents concerning Defendants' monitoring, tracking, and knowledge of investigations into whether Defendants' AI Models contain copyrighted content, including Times Content.

**REQUEST FOR PRODUCTION NO. 50:**

Documents concerning Defendants' abilities and efforts to track The Times's use of Defendants' Generative AI Products and Services.

**REQUEST FOR PRODUCTION NO. 51:**

Documents concerning the behavior of users of Defendants' Generative AI Products and Services, including (i) documents concerning the extent to which users of Defendants' Generative AI Products and Services navigate away from Defendants' Generative AI Products and Services to other websites, and (ii) documents concerning Defendants' ability and efforts to modify the

frequency with which users navigate away from Defendants’ Generative AI Products and Services to other websites.

**REQUEST FOR PRODUCTION NO. 52:**

Documents concerning the extent to which Defendants’ Generative AI Products and Services direct users to Times Content in response to user queries, including by providing outputs that directly quote Times Content, contain a URL that is intended to direct the user to Times Content, or reference The Times or one of its contributors, authors, contractors, or employees.

**REQUEST FOR PRODUCTION NO. 53:**

Documents concerning the decision to suspend Browse with Bing and whether or not to suspend Bing Chat.

**REQUEST FOR PRODUCTION NO. 54:**

Documents concerning any public statements made by You or Your employees relating to this litigation, including: (i) Your claim that the New York Times “hacked” Defendants’ Generative AI Products,<sup>2</sup> ii) Your contention that Times Content is “not significant for the model’s intended learning” and that it did not “meaningfully contribute to the training of [Your] existing models,”<sup>3</sup> iii) any contention by You that You can modify Defendants’ Text Generation AI Models to minimize the “regurgitation” of copyrighted content,<sup>4</sup> and iv) Your efforts to use artificial intelligence to “monetize news and other published content.”<sup>5</sup>

---

<sup>2</sup> See Dkt. 52 at 2, 12.

<sup>3</sup> OpenAI, *OpenAI and Journalism*, OPENAI (Jan. 8, 2024), <https://openai.com/blog/openai-and-journalism>.

<sup>4</sup> See Ryan Browne & MacKenzie Sigalos, *OpenAI CEO Sam Altman Says ChatGPT Doesn’t Need New York Times Data Amid Lawsuit*, CNBC (Jan. 18, 2024), <https://cnbc.com/2024/01/18/openai-ceo-on-nyt-lawsuit-ai-models-dont-need-publishers-data-.html#:~:text=Altman's%20comments%20echo%20remarks%20the,news%20content%20in%20the%20future>.

<sup>5</sup> See *id.*



**REQUEST FOR PRODUCTION NO. 55:**

Documents concerning any firings, terminations, lay-offs, impositions of mandatory leave, disciplinary actions, departures, resignations, or other personnel-related actions Defendants have taken or considered concerning individuals or teams involved with any ethics, harm reduction, safety, lawfulness, and responsible AI initiatives, including: (i) the removal and reinstatement of Sam Altman, (ii) the firing of Leopold Aschenbrenner and Pavel Izmailov, (iii) the appointment and role of Dee Templeton, (iv) the firing of Helen Toner, (v) the departures of Ilya Sutskever, Jan Leike, and Daniel Kokotajlo, and (vi) Microsoft's involvement in hiring and firing employees of any OpenAI entity.

**REQUEST FOR PRODUCTION NO. 56:**

Documents concerning any deliberation or decision(s) by Defendants not to publicly release research, papers, or studies related to Defendants' Generative AI Models and Defendants' Generative AI Products and Services, including any communications regarding these decisions and concerns regarding the effects of releasing this data on safety, ethics, profitability, competition, and legal obligations and/or ramifications.

**REQUEST FOR PRODUCTION NO. 57:**

Documents related to the use of Defendants' Generative AI Products and Services, including by third parties, to promulgate misinformation, disinformation, or poor quality "pink-slime" journalism, including the use of Your APIs to create websites or social media posts containing AI-generated content.

**REQUEST FOR PRODUCTION NO. 58:**

Documents concerning Your review and approval process for 1) custom GPTs made available in Your Custom GPT Store, and 2) third-party ChatGPT plug-ins made available in Your

ChatGPT plug-in store, including custom GPTs and plug-ins designed to remove paywalls, provide Journalism, or that otherwise reference The Times or Journalism.

**REQUEST FOR PRODUCTION NO. 59:**

Documents concerning Your knowledge of third-party websites that contain copies of The Times's content and Your use of GPTBot and any other web crawlers, bots, spiders, user agents, and related tools to access those third-party websites.

**REQUEST FOR PRODUCTION NO. 60:**

Documents concerning Defendants' licensing or use of Bing's search index.

**REQUEST FOR PRODUCTION NO. 61:**

Documents concerning any indemnification obligations, common interest agreements, or joint defense agreements between Defendants related to Defendants' Generative AI Products or Services.

**REQUEST FOR PRODUCTION NO. 62:**

Documents and communications concerning negotiations between Defendants and owners of copyrighted content, including The Times, over Defendants' use of copyrighted content in Defendants' Generative AI Models and Defendants' Generative AI Products or Services.

**REQUEST FOR PRODUCTION NO. 63:**

Documents concerning Your communications with other LLM creators and developers of Generative AI Products and Services regarding the use of copyrighted content to train and operate Generative AI Models and Generative AI Products and Services.

**REQUEST FOR PRODUCTION NO. 64:**

Documents concerning competition and comparisons between Defendants' Generative AI Products and Services and those of other companies, including: (i) how Generative AI tools enable You to compete in search, (ii) how Defendants' use of copyrighted materials can help You compete

with Google, (iii) any analyses and projections of how increasing Your search market share will boost Your revenue, (iv) the benefits of being a search or knowledge “destination,” and (v) Google’s reduction of search referral traffic (via SGE and direct answers) and stated goal of keeping users on its site.

**REQUEST FOR PRODUCTION NO. 65:**

Documents concerning Your public relations strategies as they pertain to Generative AI Models, Generative AI Products and Services, or copyright.

**REQUEST FOR PRODUCTION NO. 66:**

Documents concerning fundraising efforts related to Defendants’ Generative AI Products and Services, including any investor solicitation efforts and any solicitation of charitable or in-kind contributions.

**REQUEST FOR PRODUCTION NO. 67:**

Documents sufficient to show (i) projected revenue streams from Defendants’ Generative AI Products, (ii) past, present, and anticipated future subscriber numbers, and (iii) product-level revenue data from 2021 to present.

**REQUEST FOR PRODUCTION NO. 68:**

Documents concerning business development plans, strategic roadmaps, and product development plans related to (i) Defendants’ Generative AI Products and Services and (ii) any future products or services in the categories of news, cooking, consumer reviews, and sports.

**REQUEST FOR PRODUCTION NO. 69:**

Documents concerning Your policies or practices regarding public access to Your governing documents, financial statements, and conflict of interest rules.<sup>6</sup>

**REQUEST FOR PRODUCTION NO. 70:**

Documents concerning referral, affiliate, advertising, and paid search revenue generated by Defendants' Generative AI Products and Services.

**REQUEST FOR PRODUCTION NO. 71:**

Documents sufficient to show in detail, on a monthly, quarterly, or annual basis, the revenue and profit realized from Defendants' Generative AI Products and Services, including documents sufficient to show in detail, by product or service, the gross revenue generated, the costs, the gross profit margin, and the net profit margin.

**REQUEST FOR PRODUCTION NO. 72:**

Documents relating to the pricing of Defendants' Generative AI Products and Services in the United States, including price quotes, price lists, subscription amounts, price announcements, and actual prices paid by customers.

Dated: May 22, 2024

/s/ Zach Savage

Ian Crosby (*pro hac vice*)  
Genevieve Vose Wallace (*pro hac vice*)  
Katherine M. Peaslee (*pro hac vice*)  
SUSMAN GODFREY L.L.P.  
401 Union Street, Suite 3000  
Seattle, WA 98101  
Telephone: (206) 516-3880  
Facsimile: (206) 516-3883  
icrosby@susmangodfrey.com

---

<sup>6</sup> Paresh Dave, *OpenAI Quietly Scrapped a Promise to Disclose Key Documents to the Public*, WIRED (Jan. 24, 2024), <https://www.wired.com/story/openai-scrapped-promise-disclose-key-documents/>.

gwallace@susmangodfrey.com  
kpeaslee@susmangodfrey.com

Davida Brook (*pro hac vice*)  
Emily K. Cronin (*pro hac vice*)  
Ellie Dupler (*pro hac vice*)  
SUSMAN GODFREY L.L.P.  
1900 Ave of the Stars, Suite 1400  
Los Angeles, CA 90067  
Telephone: (310) 789-3100  
Facsimile: (310) 789-3150  
dbrook@susmangodfrey.com  
ecronin@susmangodfrey.com  
edupler@susmangodfrey.com

Elisha Barron (5036850)  
Zachary B. Savage (ZS2668)  
Tamar Lusztig (5125174)  
Alexander Frawley (5564539)  
Eudokia Spanos (5021381)  
SUSMAN GODFREY L.L.P.  
One Manhattan West  
New York, NY 10001  
Telephone: (212) 336-8330  
Facsimile: (212) 336-8340  
ebarron@susmangodfrey.com  
zsavage@susmangodfrey.com  
tlusztig@susmangodfrey.com  
afrawley@susmangodfrey.com  
espanos@susmangodfrey.com

Scarlett Collings (4985602)  
SUSMAN GODFREY L.L.P.  
1000 Louisiana, Suite 5100  
Houston, TX 77002  
Telephone: (713) 651-9366  
Facsimile: (713) 654-6666  
scollings@susmangodfrey.com

Steven Lieberman (SL8687)  
Jennifer B. Maisel (5096995)  
Kristen J. Logan (*pro hac vice*)  
ROTHWELL, FIGG, ERNST & MANBECK, P.C.  
901 New York Avenue, N.W., Suite 900 East  
Washington, DC 20001  
Telephone: (202) 783-6040  
Facsimile: (202) 783 6031

slieberman@rothwellfigg.com  
jmaisel@rothwellfigg.com  
klogan@rothwellfigg.com

*Attorneys for Plaintiff*  
*The New York Times Company*

**CERTIFICATE OF SERVICE**

I declare that I am employed with the law firm of Susman Godfrey L.L.P., whose address is 1900 Avenue of the Stars, Suite 1400, Los Angeles, CA 90067. I am not a party to the within cause, and I am over the age of eighteen years.

I further declare that on May 22, 2024, I served a copy of:

**THE NEW YORK TIMES COMPANY'S SECOND SET OF REQUESTS FOR PRODUCTION TO OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, AND OPENAI HOLDINGS, LLC**

- ☒ **BY ELECTRONIC SERVICE [Fed. Rule Civ. Proc. Rule 5(b)(2)(E)]** by electronically mailing a true and correct copy through Susman Godfrey L.L.P.'s electronic mail system to the email address(es) set forth below, or as stated on the attached service list per agreement in accordance with Fed. Rule Civ. Proc. Rule 5(b)(2)(E).
- ☐ **BY PERSONAL SERVICE** I caused to be delivered such envelope by hand to the offices of the addressee.

**See Attached Service list**

I declare under penalty of perjury that the following is true and correct.

Executed in Los Angeles, California, this 22<sup>nd</sup> day of May, 2024.

Ellie Dupler  
(Type)

Ellie Dupler  
(Signature)

## SERVICE LIST

### [OpenAICopyright@mofo.com](mailto:OpenAICopyright@mofo.com)

Joseph C. Gratz  
Vera Ranieri  
Morrison & Foerster LLP  
425 Market Street  
San Francisco, CA 94105-2482  
Telephone: (415) 268-6066  
Facsimile: (415) 268-7522  
Email: jgratz@mofo.com  
Email: Vranieri@mofo.com

Allyson R. Bennett  
Rose S. Lee  
Morrison & Foerster LLP  
707 Wilshire Boulevard, Suite 6000  
Los Angeles CA 90017  
Telephone: (213) 892-5200  
Facsimile: (213) 892-5454  
Email: Abennett@mofo.com  
Email: RoseLee@mofo.com

Eric Nikolaides  
Morrison & Foerster LLP  
250 W. 55th St  
New York, NY 10019  
Telephone: 212-336-4061  
Email: enikolaides@mofo.com

#### **Attorneys for Defendants**

**OpenAI, Inc., OpenAI LP, OpenAI GP, LLC,  
OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global  
LLC, OAI Corporation, LLC, And OpenAI  
Holdings, LLC**

### [OpenAICopyrightLitigation.lwteam@lw.com](mailto:OpenAICopyrightLitigation.lwteam@lw.com)

Andrew Gass  
Joseph Richard Wetzel , Jr.,  
LATHAM & WATKINS, LLP  
505 Montgomery Street, Ste. 2000  
San Francisco, CA 94111  
Telephone: (415)391-0600  
Facsimile: (415)-395-8095  
Email: andrew.gass@lw.com  
Email: joe.wetzel@lw.com

Allison Levine Stillman  
Luke Budiardjo  
LATHAM & WATKINS LLP  
1271 Avenue of the Americas  
New York, NY 10020  
Telephone: (212)906-1200  
Facsimile: 212-751-4864  
Email: alli.stillman@lw.com  
Email: luke.budiardjo@lw.com

Sarang Damle  
LATHAM & WATKINS LLP  
555 Eleventh Street NW, Suite 1000  
Washington, DC 20004  
Telephone: (202)637-2200  
Facsimile: 202-637-2201  
Email: sy.damle@lw.com

#### **Attorneys for Defendants**

**OpenAI, Inc., OpenAI LP, OpenAI GP, LLC,  
OpenAI, LLC, OpenAI OpCo LLC, OpenAI  
Global LLC, OAI Corporation, LLC, And  
OpenAI Holdings, LLC**



[kvpoi@keker.com](mailto:kvpoi@keker.com)

Robert A. Van Nest  
Paven Malhotra  
R. James Slaughter  
Katie Lynn Joyce  
Michelle S. Ybarra  
Nicholas S Goldberg  
Sarah Salomon  
Thomas Edward Gorman  
KEKER, VAN NEST & PETERS LLP  
633 Battery Street  
San Francisco, CA 94111  
Telephone: (415) 391-5400  
Facsimile: (415) 397-7188  
Email: rvannest@keker.com  
Email: pmalhotra@keker.com  
Email: rslaughter@keker.com  
Email: kjoyce@keker.com  
Email: mybarra@keker.com  
Email: ngoldberg@keker.com  
Email: ssalomon@keker.com  
Email: tgorman@keker.com

**Attorneys for Defendants  
OpenAI, Inc., OpenAI LP, OpenAI GP, LLC,  
OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global  
LLC, OAI Corporation, LLC, And OpenAI  
Holdings, LLC**

[NewYorkTimes\\_Microsoft\\_OHS@orrick.com](mailto:NewYorkTimes_Microsoft_OHS@orrick.com)

Annette L. Hurst  
ORRICK, HERRINGTON & SUTCLIFFE LLP  
The Orrick Building  
405 Howard Street  
San Francisco, CA 94105  
Telephone: (415)773-5700  
Facsimile: (415)773-5759  
Email: ahurst@orrick.com

Christopher J. Cariello  
ORRICK, HERRINGTON & SUTCLIFFE LLP  
51 West 52nd Street  
New York, NY 10019  
Telephone: (212) 506-3778  
Facsimile: (212) 506-5151  
Email: ccariello@orrick.com

Sheryl Koval Garko  
Laura Brooks Najemy  
ORRICK, HERRINGTON & SUTCLIFFE LLP  
222 Berkeley Street, Suite 2000  
Boston, MA 02116  
Telephone: (617) 880-1800  
Facsimile: (617) 8801-1801  
Email: sgarko@orrick.com  
Email: lnajemy@orrick.com

**Attorneys for Defendant  
Microsoft Corporation**

[MicrosoftNYClassActionFDBR@faegredrinker.com](mailto:MicrosoftNYClassActionFDBR@faegredrinker.com)

Jeffrey S. Jacobson  
FAEGRE DRINKER BIDDLE & REATH LLP  
1177 Avenue of the Americas  
New York, New York 10036  
Telephone: (212) 248-3191  
Email: jeffrey.jacobson@faegredrinker.com

Jared B. Briant  
FAEGRE DRINKER BIDDLE & REATH LLP  
1144 15th Street, Suite 3400  
Denver, Colorado 80202  
Telephone (303) 607-3588  
Email: jared.briant@faegredrinker.com

**Attorneys for Defendant  
Microsoft Corporation**